J. Child Lang. **41** (S1) (2014), 124–131. © Cambridge University Press 2014
doi:10.1017/S0305000914000142

What we have learned*

BRIAN MACWHINNEY
Department of Psychology, Carnegie Mellon University

This special issue celebrates the fortieth anniversary of the *Journal of Child Language* by presenting reflections on the past and future of research on child language acquisition. In the spirit of this celebration, I review informally the major principles we have identified, the ideas we have rejected, and some of the crucial issues remaining to be explored. Beginning with the positive findings, we can sort the principles we have confirmed into four categories: learning principles, processing principles, neurolinguistic principles, and social interaction principles.

Let us begin with consideration of some of the overarching principles governing child language learning.

1. *Structure and productivity*. Wundt, Leopold, Brown, Berko, and many others have demonstrated the pervasive impact of structure building in language acquisition. Children build structures and generalizations on all linguistic levels, including phonology, morphology, semantics, syntax, and conversation.
2. *Compositionality*. In general, children learn simple structures before more complex structures. For example, they first learn words and then learn to put these words together into phrases. In phonology, children learn to control their vocal cords and tongue before producing consistent target sounds (Oller, 2000). At the same time, children can also pick up complex patterns as single rote items from which they then pull out smaller components over time.
3. *Generalization*. Structures arise from generalization over instances. Within each of the linguistic levels, generalization takes on slightly different forms. In phonology, phonemic categories are abstracted as central tendencies across exemplars. In word learning, concepts are extracted from specific examples.
4. *Lexical basis of syntax*. In syntax, lexical combinations give rise to item-based patterns and then feature-based patterns through a process of increasing levels of generalization (MacWhinney, in press). The pervasive nature of this process is reflected in the tight correlation between lexicon at 1;8 and grammatical abilities eight months later (Bates & Goodman, 1999).

[*] e-mail: macw@cmu.edu

https://doi.org/10.1017/S0305000914000142 Published online by Cambridge University Press

5. *Statistical learning*. Statistical learning operates sequentially to extract candidate words and distributionally to extract sound categories and phonotactic patterns (Thiessen & Erickson, 2014).
6. *Bootstrapping*. The categories and patterns that arise from generalization and statistical learning can then facilitate acquisition of further categories and patterns. For example, learning of Hungarian vowel contrasts then allows the child to acquire Hungarian vowel harmony. There is an interplay between generalization and episodic encoding of specific instances. Some children move quickly to form generalizations; others are more conservative. Some encode episodes in great detail; others use less precise encodings.
7. *Availability*. Initially, children extract patterns from forms that are easily detectable and highly frequent. Detectability is facilitated when cues appear at the end of words and phrases or when they are constructed as full syllables or separable properties (Slobin, 1985).
8. *Prototypicality*. Because patterns are extracted from specific instances, early patterns function as prototypes. For example, Cantonese-speaking children work initially with a sentence pattern in which animacy, topicality, and word order are all expressed through the subject (Chan, Lieven & Tomasello, 2009). Prototype organization is also evident in phonology (Kuhl, 1991) and semantics (Mervis & Rosch, 1981).
9. *Reliability*. Although availability, detectability, and prototypicality play major roles in early learning, later learning is influenced more by cue or pattern reliability (McDonald & MacWhinney, 1991).
10. *Ambiguity*. Because ambiguity represents a specific form of unreliability, it can delay acquisition. For example, the use of the German article *der* to mark either masculine nominative or feminine/plural genitive can lead to occasional confusions (Pelham, 2011).
11. *Distinctions are functionally contrastive*. Children learn to set up categories and distinctions on the basis of contrasts in meaning or function (Clark, 1987). Some distinctions map directly to contrasts in the real world, others to obligatory grammatical distinctions, such as gender or verb class. The extraction of distinctions arises primarily through self-organization.

Many of the core principles derived from the study of adult language processing apply with equal force to children's processing and learning.

1. *Competition*. Competition arises between patterns that have been formulated on different levels of generality or based on conflicting datasets.
2. *Multilingual competition*. We have increasing evidence for concurrent activation of multiple languages during processing in adults. Although adults develop attentional processes that can limit crosstalk, there is

https://doi.org/10.1017/S0305000914000142 Published online by Cambridge University Press

still interlanguage transfer, particularly in the context of code-switching environments. These same transfer effects can be identified in children (Yip & Matthews, 2007).

3. *Combination vs. rote*. The most common competition is between analytic or combinatorial forms and synthetic or rote forms, although there are also competitions between alternative analytic forms and analogistic patterns (MacWhinney, 1978).
4. *Minimization*. We have learned that the processor operates in ways that minimize attachment, distance, and memory costs (O'Grady, 2005). These forces operate in children as well as adults, but they can be seen most clearly in older children.
5. *Formulation*. Like adults, children correct their errors, both for form and substance (MacWhinney & Osser, 1977).

Ongoing advances in neuroscience continue to teach us about the neuronal bases for language development.

1. *Early differentiation*. Genetic and epigenetic variation can lead to different patterns of brain structure and connectivity, some of which lead to language disorders (Bishop, 2013).
2. *Language sparing*. Children who experience brain lesions often develop normal language functioning (MacWhinney, Feldman, Sacco & Valdes-Perez, 2000). Language seems to be uniquely favored in this regard, because of the fact that it utilizes alternative methods for constructing messages.
3. *Vulnerability of timing*. Problems with the timing of neuronal integration seem to be central to various language disorders, as well as delays in normal acquisition (Howell, 2011; Roelofs, 2011).
4. *Articulatory challenges*. The control of phonological output is a major development challenge. It involves a high level of precise control and coordination of fine-grained motions that must be extended to a wide variety of articulatory targets (Browman & Goldstein, 1992).
5. *Entrenchment*. The development of prototypes is one reflection of the general tendency for neuronal patterns to become entrenched through learning. This effect is the major determinant underlying age-related declines in second language learning ability (MacWhinney, 2012).
6. *Plasticity*. Perhaps the most dramatic evidence for plasticity is the fact that ASL speakers make use of basic language processing areas for controlling the production of sign language (Corina, Vaid & Bellugi, 1992).
7. *Specialization*. Although neurolinguistic development reflects flexibility, sparing, and plasticity, there are also limits to the plasticity of certain cognitive functions. For example, although children with focal lesions recover language functions, they cannot recover from

https://doi.org/10.1017/S0305000914000142 Published online by Cambridge University Press

right hemiplegia. These limitations are particularly evident in the entrenchment of motor and auditory processes (Guion, Flege, Akahane-Yamada & Pruitt, 2000).

8. *Individual variation*. Children differ markedly during the earliest stages of learning. Some emphasize clarity or faithfulness in pronunciation; others use filler words (Peters & Strömqvist, 1996), grunts, and mushy articulations. Much of this variation is based on underlying genetic variation, as further filtered through epigenesis. Despite this wide initial variation, all normal children end up learning their native languages.

We have also learned a great deal about the impact of social structures on children's language learning.

1. *Trust and identification*. Children operate within a zone of proximal development grounded on trust in and identification with their caretakers (Tomasello, 2003).
2. *Common ground*. Like adults, children rely on the construction of common ground to advance conversation (E. Clark, 2014).
3. *Cooperativeness*. Caretakers try to facilitate children's learning of word meanings by giving them labels that correspond to what they think the children have in mind (Brown, 1958).
4. *Simple input*. Caretakers simplify the child's learning task by providing simple input, often in the shape of single-word utterances or labels (MacWhinney, in press).
5. *Scaffolding*. Caretakers provide further support for language learning by scaffolding utterances and exchanges through completion, recasting, and repetition (Bohannon & Stanowicz, 1988).
6. *Fine-tuning*. Caretakers provide input that is finely tuned to the child's advancing developmental level (Sokolov, 1993).
7. *Input variability*. Some children receive extensive rich input; others little input at all (Hart & Risley, 1995). Children who receive less adult input will have smaller vocabularies and subsequently weaker academic performance (Snow, 1999).
8. *Code selection*. Multilingual children show early sensitivity to the many social conditions determining code selection (Genesee, Boivin & Nicolaidis, 1996).

It would be easy to expand each of these lists, by including additional topics and more specific instantiations of each of the principles. My hope is that a simple listing of this type will make it clear how much progress we have made in the last half-century of child language research. Of course, not all of these findings are universally accepted. It is important for academics to challenge ideas through debate and discussion, and each of these principles

https://doi.org/10.1017/S0305000914000142 Published online by Cambridge University Press

will surely be further refined through this process. However, we can see from this listing that we now have an increasingly solid grasp on the basic nature of first language acquisition.

Much of this progress has occurred through studies challenging earlier simple formulations. Let us consider some of the ideas that have been rejected in this way.

1. *Degenerate input*. According to Chomsky (1980), the input to the child is so full of retracings, errors, and incompletions that it provides no clear indication of the possible sentences of the language. However, Newport, Gleitman, and Gleitman (1977) showed that "the speech of mothers to children is unswervingly well-formed".
2. *The logical problem*. Gold (1967) demonstrated that, under certain assumptions, language identification in the limit would be unattainable. However, more recent work (A. Clark, 2014; MacWhinney, 2004) has questioned the relevance of Gold's result.
3. *Universal phonology*. Early accounts of the development of phonology claimed that, during the babbling period, children produced all the possible sounds of human language. However, closer studies have shown that the range of sounds produced during babbling is limited to certain basic patterns (Cruttenden, 1970; Davis & MacNeilage, 1995).
4. *Discontinuities*. Views of language development that emphasize discontinuities, such as Jakobson's (1968) claims about the discontinuity between babbling and early phonology, have been rejected.
5. *Shaping*. Behaviorist accounts of language learning were based on notions of shaping and differential reinforcement that have now been discarded (Chomsky, 1959) .
6. *Cognitive determination*. During the 1970s, researchers pursued the idea that advances in object permanence during stage 6 of the sensorimotor period would trigger advances in vocabulary and syntax. However, studies failed to confirm these predictions (Corrigan, 1979).
7. *Mutual exclusivity*. The idea that children have a built-in constraint against referring to things by different names has been invalidated by evidence that bilingual children are happy to accept multiple names, as well as children's willingness to learn forms that describe things in alternative contexts.
8. *Parameters*. The idea that children learn language by setting values on a small set of syntactic parameters was difficult to formulate and never received clear empirical support.
9. *Back propagation*. Attempts to model language acquisition using back propagation networks without additional computational structures failed to account for important aspects of the learning of morphology and syntax.

https://doi.org/10.1017/S0305000914000142 Published online by Cambridge University Press

Some of these concepts may be resurrected in modified form in future accounts, but for now they play a marginal role.

Child language studies are being infused by many new cross-disciplinary insights and methodological developments. Laboratory methods such as event-related potential (ERP), near-infrared spectroscopy (NIRS), preferential looking, and eye-tracking provide us with increasing detail regarding the flow of processing. The development of standard stimulus sets for objects, words, sounds, and interactions has added further structure to these experimental efforts, tracking infants' abilities on a month-by-month basis as they progress through language learning and socialization, although often through a cross-sectional methodology. Child language studies are also beginning to link up with work in developmental biology that characterizes the genetic and epigenetic basis of language development. Perhaps the most dramatic methodological advances have come from the application of computational technology to child language. Using corpora from CHILDES, researchers are formulating increasingly detailed computational models of the learning of phonology, lexicon, syntax, and conversation.

Despite all of this progress and promise, there are large areas of language learning that are not receiving adequate attention. Although we have formulated many possible accounts and mechanisms, our understanding of the causes of disorders such as stuttering and specific language impairment remains vague. Another major gap involves our understanding of how children refine word meanings over time across episodes in real life, using feedback from multiple encounters. By restricting our attention to laboratory experiments, we have failed to develop a picture of how children sharpen word shapes, and how they come to deal with polysemy, homonymy, and ambiguity. This problem is particularly acute when we come to the learning of school-based concepts such as SYMMETRY, DEPENDABILITY, MAJORITY, SPECIES, and POLITICS (MacWhinney, 2005). Understanding the refinement of such terms will require us to dig more deeply into children's learning of social and scientific structures. To do this, we will need to link up work on the structuring of mental models of complex systems and processes in adults with their development in children.

There is a similar gap in our understanding of the development of persuasion. When I began recording my sons' language development, it was my intention to study this, but other things have intervened. Fortunately, Karniol (2010) has shown how one can use transcripts from CHILDES and other sources to make progress on this issue.

These last two gaps illustrate an overarching problem we have in terms of tracking acquisitional patterns that operate across longer timeframes. Relying only on cross-sectional data, we cannot understand the dynamic aspects of these processes. For example, we need to understand exactly what interactions and social configurations can lead a bilingual child to reject use

https://doi.org/10.1017/S0305000914000142 Published online by Cambridge University Press

of the home language, when entering school. To study such processes, we must commit ourselves to increasingly ambitious attempts to record the development of individual children or perhaps several children in similar sociolinguistic contexts. The growth of methods for collecting, transcribing, and analyzing large video corpora may make this increasingly possible in the future (Roy, 2013).

REFERENCES

Bates, Elizabeth & Goodman, Judith (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (ed.), *The emergence of language*, 29–80. Mahwah, NJ: Lawrence Erlbaum Associates.

Bishop, Dorothy (2013). Developmental cognitive genetics: how psychology can inform genetics and vice versa. *Quarterly Journal of Experimental Psychology* **59**, 1153–68.

Bohannon, N. & Stanowicz, L. (1988). The issue of negative evidence: adult responses to children's language errors. *Developmental Psychology* **24**, 684–9.

Browman, C. P. & Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica* **49**, 155–80.

Brown, Roger (1958). How shall a thing be called? *Psychological Review* **65**, 14–21.

Chan, Angel, Lieven, Elena & Tomasello, Michael (2009). Children's understanding of the agent–patient relations in the transitive construction: cross-linguistic comparisons between Cantonese, German and English. *Cognitive Linguistics* **20**, 267–300.

Chomsky, Noam (1959). Review of Skinner's "Verbal Behavior". *Language* **35**, 26–58.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Clark, Alexander (2014). Learnability. In B. MacWhinnney & W. O'Grady (eds), *Handbook of language emergence*. New York: Wiley.

Clark, Eve (1987). The Principle of Contrast: a constraint on language acquisition. In B. MacWhinney (ed.), *Mechanisms of language acquisition*, 1–34. Hillsdale, NJ: Lawrence Erlbaum Associates.

Clark, Eve (2014). Common ground. In B. MacWhinney & W. O'Grady (eds), *Handbook of language emergence*. New York: Wiley.

Corina, David P., Vaid, Jyotsna & Bellugi, Ursula (1992). The linguistic basis of left hemisphere specialization. *Science* **255**(5049), 1258–60.

Corrigan, R. (1979). Cognitive correlates of language: differential criteria yield differential results. *Child Development* **50**, 617–31.

Cruttenden, A. (1970). A phonetic study of babbling. *British Journal of Disorders of Communication* **5**, 110–7.

Davis, B. L. & MacNeilage, Peter F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research* **38**, 1199–211.

Genesee, F., Boivin, I. & Nicolaidis, E. (1996). Talking with strangers: a study of bilingual children's communicative competence. *Applied Psycholinguistics* **17**, 427–42.

Gold, E. M. (1967). Language identification in the limit. *Information and Control,* **10**, 447–474.

Guion, S. G., Flege, J. E., Akahane-Yamada, R. & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: the case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America* **107**, 2711–24.

Hart, B. & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Paul H. Brookes.

Howell, Peter (2011). *Recovering from stuttering*. New York: Psychology Press.

Jakobson, R. (1968). *Child language, aphasia and phonological universals*. The Hague: Mouton.

Karniol, Rachel (2010). *Social development as preference management*. New York: Cambridge University Press.

https://doi.org/10.1017/S0305000914000142 Published online by Cambridge University Press

Kuhl, Patricia (1991). Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* **50**, 93–107.
MacWhinney, Brian (1978). The acquisition of morphophonology. *Monographs of the Society for Research in Child Development* **43**(1), 1–123.
MacWhinney, Brian (2004). A multiple process solution to the logical problem of language acquisition. *Journal of Child Language* **31**, 883–914.
MacWhinney, Brian (2005). Can our experiments illuminate reality? In L. Gershkoff-Stowe & D. Rakison (eds), *Building object categories in developmental time*, 301–8. Mahwah, NJ: Lawrence Erlbaum Associates.
MacWhinney, Brian (2012). The logic of the Unified Model. In S. Gass & A. Mackey (eds), *The Routledge handbook of second language acquisition*, 211–27. New York: Routledge.
MacWhinney, Brian (in press). Item-based patterns in early syntactic development. In T. Herbst (ed.), *Valency relations*. Berlin: Springer.
MacWhinney, Brian, Feldman, Heidi, Sacco, Kelly & Valdes-Perez, Raul (2000). Online measures of basic language skills in children with early focal brain lesions. *Brain and Language* **71**, 400–31.
MacWhinney, Brian & Osser, H. (1977). Verbal planning functions in children's speech. *Child Development* **48**, 978–85.
McDonald, Janet & MacWhinney, Brian (1991). Levels of learning: a microdevelopmental study of concept formation. *Journal of Memory and Language* **30**, 407–30.
Mervis, C. & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology* **32**, 89–115.
Newport, E, Gleitman, H. & Gleitman, L. (1977). Mother, I'd rather do it myself: some effects and noneffects of maternal speech style. In C. Snow (Ed.), *Talking to children: language input and acquisition*. Cambridge: Cambridge University Press.
O'Grady, William (2005). *Syntactic carpentry*. Mahwah, NJ: Lawrence Erlbaum Associates.
Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Lawrence Erlbaum Associates.
Pelham, Sabra (2011). The input ambiguity hypothesis and case blindness: an account of cross-linguistic and intra-linguistic differences in case errors. *Journal of Child Language* **38**, 235–72.
Peters, A. M. & Strömqvist, S. (1996). The role of prosody in the acquisition of grammatical morphemes. In J. Morgan & K. Demuth (eds), *From signal to syntax*, 215–32. Hillsdale, NJ: Lawrence Erlbaum.
Roelofs, Ardi (2011). Modeling the attentional control of vocal utterances: from Wernicke to WEAVER+. In J. Guendozi, F. Loncke & M. Williams (eds), *The handbook of psycholinguistic and cognitive processes: perspectives in communication disorders*, 189–208. New York: Psychology Press.
Roy, Brandon. (2013). *The birth of a word*. (Unpublished PhD thesis) MIT.
Slobin, D. (1985). Crosslinguistic evidence for the language-making capacity. In D. Slobin (ed.), *The crosslinguistic study of language acquisition. Volume 2: theoretical issues*, 1157–256. Hillsdale, NJ: Lawrence Erlbaum.
Snow, Catherine E. (1999). Social perspectives on the emergence of language. In B. MacWhinney (ed.), *The emergence of language*, 257–76. Mahwah, NJ: Lawrence Erlbaum Associates.
Sokolov, J. L. (1993). A local contingency analysis of the fine-tuning hypothesis. *Developmental Psychology* **29**, 1008–23.
Thiessen, Erik & Erickson, Lucy (2014). Perceptual development and statistical learning. In B. MacWhinney & W. O'Grady (eds), *Handbook of language emergence*. New York: Wiley.
Tomasello, Michael (2003). *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Yip, V. & Matthews, S. (2007). *The bilingual child: early development and language contact*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S0305000914000142 Published online by Cambridge University Press